

Συσχέτιση (Pearson r) και Ανάλυσης Διακύμανσης (ANOVA) με ένα παράγοντα

1. Να χρησιμοποιήσετε το αρχείο "gssnet.sav" έτσι ώστε να βρεθούν όλα τα ζεύγη συσχετίσεων ανάμεσα στην «εκπαίδευση μητέρας», εκπαίδευση πατέρα», «εκπαίδευση συμμετέχοντα» και «εκπαίδευση συζύγου» (μεταβλητές *maeduc*, *paeduc*, *educ*, και *seduc*)
 - a. Είναι κατάλληλος ο συντελεστής συσχέτισης *r* για τις συσχετίσεις των παραπάνω μεταβλητών; Εξηγήστε.
 - b. να βρεθεί η μεταβλητή που προβλέπει καλύτερα την «εκπαίδευση συμμετέχοντα» καθώς και το ποσοστό της ερμηνευόμενης διακύμανσης της «εκπαίδευσης συμμετέχοντα» από την μεταβλητή αυτή.
2. Στα πλαίσια της μελέτης των παραγόντων που μπορεί να ερμηνεύσουν την αντιλαμβανόμενη ευτυχία θα μελετηθεί η επίδραση της οικογενειακής κατάστασης (μεταβλητή *maritalb*) στην αντιλαμβανόμενη «ευτυχία» (μεταβλητή *happy*) από το αρχείο **SS_EKKE_2011.sav** που περιλαμβάνει τα δεδομένα κοινωνικής έρευνας του ΕΚΚΕ του 2011. Πριν εκτελεστεί η ανάλυση θα συγχωνευτεί η κατηγορία 2, «Νομικά αναγνωρισμένη συμβίωση (σύμφωνο συμβίωσης)» με την κατηγορία 1 «Γάμος (θρησκευτικός ή πολιτικός)» αφού στην κατηγορία ανήκει μόλις το 0,9% των ερωτηθέντων. Είναι επίσης κατανοητή η συγγένεια των δυο κατηγοριών. Για το σκοπό αυτό θα χρησιμοποιηθεί η ρουτίνα του SPSS για να δημιουργηθεί η νέα μεταβλητή της οικογενειακής κατάστασης *maritalb_n*:

```
RECODE maritalb (2=1) (ELSE=Copy) INTO maritalb_n .
```

```
EXECUTE.
```

Για να διευκολυνθεί η κατανόηση των αποτελεσμάτων θα αντιγραφεί το περιεχόμενο του πεδίου Values της μεταβλητής *marital* στην αντίστοιχη θέση της μεταβλητής *maritalb_n*. Θα επιλεγούν για την ανάλυση άτομα ηλικίας >30 ετών (μεταβλητή *agea*) ελαχιστοποιώντας το συγχυτικό ρόλο της ηλικίας στο αποτέλεσμα καθώς είναι γνωστό ότι τόσο η οικογενειακή κατάσταση όσο και η ευτυχία συνδέονται σημαντικά με την ηλικία.

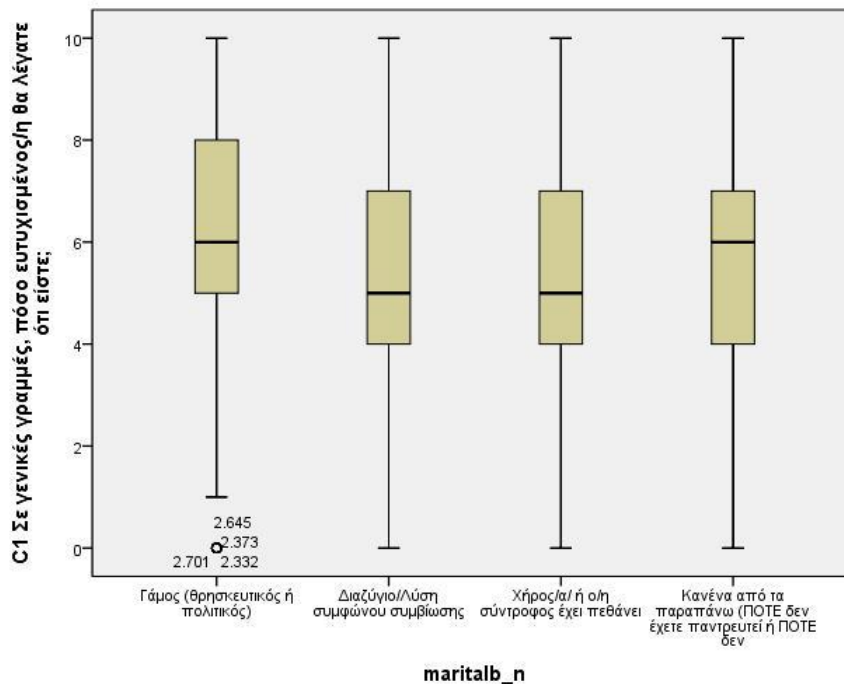
- a. Να δημιουργηθούν θηκογράμματα της «Ευτυχίας» ανά κατηγορία «οικογενειακής κατάστασης». Με βάση τα θηκογράμματα υπάρχουν οι προϋποθέσεις εκτέλεσης της Ανάλυσης Διακύμανσης;
- b. Να γίνει έλεγχος της ισότητας των διακυμάνσεων της «ευτυχίας» στις 4 κατηγορίες «οικογενειακής κατάστασης» (θα εκτελεστεί η ανάλυση διακύμανσης και στη επιλογή "Options" θα μπει «τικ» στην επιλογή "homogeneity of variance test". Το αποτέλεσμα αυτό συμφωνεί με εκείνο του ερωτήματος a);
- c. Ας υποθέσουμε ότι ικανοποιούνται οι προϋποθέσεις εκτέλεσης της Ανάλυσης Διακύμανσης. Μπορεί να απορριφθεί η μηδενική υπόθεση ότι η μέση ευτυχία είναι ίδια σ' όλες τις κατηγορίες οικογενειακής κατάστασης;
- d. Ποιες κατηγορίες οικογενειακής κατάστασης διαφέρουν σημαντικά μεταξύ τους χρησιμοποιώντας τον έλεγχο Bonferroni;
- e. Γράψτε μια σύντομη αναφορά σχολιάζοντας το αποτέλεσμα

Απάντηση

Μετά την δημιουργία της μεταβλητής “maritalb_n” από την επιλογή “data=>select cases” και τρόπο επιλογής “if condition is satisfied”, δίνουμε τη συνθήκη “agea >30”. Επίσης χρειάζεται στο “data view” να δηλωθεί η τιμή 88 της “maritalb_n” ως missing. Η κατανομή συχνοτήτων (Analyze=>frequencies) πρέπει να δημιουργεί τον παρακάτω πίνακα

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 Γάμος (θρησκευτικός ή πολιτικός)	1389	65,0	65,5	65,5
	4,00 Διαζύγιο/Λύση συμφώνου συμβίωσης	139	6,5	6,6	72,1
	5,00 Χήρος/α/ ή ο/η σύντροφος έχει πεθάνει	294	13,8	13,9	86,0
	6,00 Κανένα από τα παραπάνω (ΠΟΤΕ δεν έχετε παντρευτεί ή ΠΟΤΕ δεν	297	13,9	14,0	100,0
	Total	2119	99,1	100,0	
Missing	88,00	19	,9		
Total		2138	100,0		

- a. Από το θηκόγραμμα που παρουσιάζεται στο παρακάτω γράφημα δεν προκύπτει ισχυρή απόκλιση από την κανονική κατανομή ούτε παρουσία ακραίων τιμών σε καμιά από τις 4 κατηγορίες οικογενειακής κατάστασης και ως εκ τούτου είναι έγκυρο το αποτέλεσμα της Ανάλυσης Διακύμανσης. Εξάλλου τα μεγάλα δείγματα που παρατηρούνται στην περίπτωσή μας απαλείφουν την προϋπόθεση της κανονικότητας. Συμπληρωματικά με βάση το ενδοτεαρτημοριακό εύρος δεν παρατηρούνται διαφορές ως προς την μεταβλητότητα μεταξύ των 4 κατηγοριών.



- b. Από το αποτέλεσμα του ελέγχου Levene που δίνεται στον παρακάτω πίνακα δεν προκύπτουν σημαντικές διαφορές ως προς την διακύμανση των τεσσάρων πληθυσμών οικογενειακής κατάστασης $Levene(3,2086)=1,22, p=0,301$). Το αποτέλεσμα συμφωνεί με εκείνο του ερωτήματος α.

Test of Homogeneity of Variances

happy C1 Σε γενικές γραμμές, πόσο ευτυχισμένος/η θα λέγατε ότι είστε;

Levene Statistic	df1	df2	Sig.
1,219	3	2086	,301

- c. Από την εκτέλεση την Ανάλυσης διακύμανσης με ένα παράγοντα και το γενικό έλεγχο με την κατανομή F που δίνεται παρακάτω, απορρίπτεται η μηδενική υπόθεση ότι η μέση ευτυχία είναι ίδια σ' όλες τις κατηγορίες οικογενειακής κατάστασης ($F(3,2086)=25,62, p<0,001$);

ANOVA

happy C1 Σε γενικές γραμμές, πόσο ευτυχισμένος/η θα λέγατε ότι είστε;

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	346,978	3	115,659	25,616	,000
Within Groups	9418,611	2086	4,515		
Total	9765,589	2089			

- d. Από τον παρακάτω πίνακα με τις συγκρίσεις μέσω των τιμών ανά δυο για τις 4 κατηγορίες οικογενειακής Κατάστασης με τη μέθοδο bonfferoni προκύπτει σημαντικά υψηλότερο μέσο επίπεδο ευτυχίας της κατηγορίας Γάμος (θρησκευτικός ή πολιτικός) σε σύγκριση με όλες τις υπόλοιπες κατηγορίες. Η κατηγορία «Κανένα από τα παραπάνω έχετε παντρευτεί ή ΠΟΤΕ» δηλώνει σημαντικά υψηλότερο μέσο επίπεδο ευτυχίας από την κατηγορία « Χήρος/α/ ή ο/η σύντροφος έχει πεθάνει».

Multiple Comparisons

Dependent Variable: happy C1 Σε γενικές γραμμές, πόσο ευτυχισμένος/η θα λέγατε ότι είστε;
Bonferroni

(I) maritalb_n	(J) maritalb_n	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1,00 Γάμος (θρησκευτικός ή πολιτικός)	4,00 Διαζύγιο/Λύση συμφώνου συμβίωσης	,935*	,190	,000	,43	1,44
	5,00 Χήρος/α/ ή ο/η σύντροφος έχει πεθάνει	1,029*	,138	,000	,67	1,39
	6,00 Κανένα από τα παραπάνω (ΠΟΤΕ δεν έχετε παντρευτεί ή ΠΟΤΕ δεν	,530*	,137	,001	,17	,89
4,00 Διαζύγιο/Λύση συμφώνου συμβίωσης	1,00 Γάμος (θρησκευτικός ή πολιτικός)	-,935*	,190	,000	-1,44	-,43
	5,00 Χήρος/α/ ή ο/η σύντροφος έχει πεθάνει	,094	,220	1,000	-,49	,68
	6,00 Κανένα από τα παραπάνω (ΠΟΤΕ δεν έχετε παντρευτεί ή ΠΟΤΕ δεν	-,405	,220	,395	-,98	,18
5,00 Χήρος/α/ ή ο/η σύντροφος έχει πεθάνει	1,00 Γάμος (θρησκευτικός ή πολιτικός)	-1,029*	,138	,000	-1,39	-,67
	4,00 Διαζύγιο/Λύση συμφώνου συμβίωσης	-,094	,220	1,000	-,68	,49
	6,00 Κανένα από τα παραπάνω (ΠΟΤΕ δεν έχετε παντρευτεί ή ΠΟΤΕ δεν	-,499*	,176	,028	-,96	-,03
6,00 Κανένα από τα παραπάνω (ΠΟΤΕ δεν έχετε παντρευτεί ή ΠΟΤΕ δεν	1,00 Γάμος (θρησκευτικός ή πολιτικός)	-,530*	,137	,001	-,89	-,17
	4,00 Διαζύγιο/Λύση συμφώνου συμβίωσης	,405	,220	,395	-,18	,98
	5,00 Χήρος/α/ ή ο/η σύντροφος έχει πεθάνει	,499*	,176	,028	,03	,96

*. The mean difference is significant at the 0.05 level.

(η παράγραφος αυτή δείχνει πώς θα γραφεί η σύγκριση των μέσων τιμών σε μια εργασία: Συνήθως οι μέσες τιμές των ομάδων παρουσιάζονται σε ένα πίνακα

(βρίσκεται παρακάτω) από την χαμηλότερη προς την υψηλότερη τιμή τοποθετώντας σύμβολα που εκφράζουν τη σημαντικότητα των διαφορών. Γίνονται $k(k-1)/2$ (k ο αριθμός κατηγοριών) συγκρίσεις οι οποίες στο θέμα μας είναι $(4*3)/2=6$ χρησιμοποιώντας τον προηγούμενο πίνακα των συγκρίσεων Bonferroni. Ξεκινώντας από την μικρότερη τιμή (5,06) την συγκρίνουμε με την επόμενη (5,15) και στη συνέχεια την συγκρίνουμε με την μεθεπόμενη (5,56) κλπ. Στη συνέχεια ξεκινάμε από τη δεύτερη μικρότερη (5,15) και την συγκρίνουμε με την επόμενη της 5,56 με την μεθεπόμενη 6,09. Αν δύο μέσες τιμές δεν διαφέρουν σημαντικά υπογραμμίζονται από το ίδιο ευθύγραμμο τμήμα. Τέλος σημειώνονται με το ίδιο γράμμα όσες μέσες τιμές ανήκουν στο ίδιο ευθύγραμμο τμήμα.

	5,00 Χήρος/α/ ή ο/η σύντροφος έχει πεθάνει	4,00 Διαζύγιο/Λύση συμφώνου συμβίωσης	6,00 Κανένα από τα παραπάνω (ΠΟΤΕ δεν έχετε παντρευτεί ή ΠΟΤΕ δεν	1,00 Γάμος (θρησκευτικός ή πολιτικός)	Total
maritalb_n					
happy	5,06 _α	5,15 _{αβ}	5,56 _β	6,09 _γ	5,81

Σημείωση. Μέσες τιμές με κοινό σύμβολο (ανήκουν στην ίδια γραμμή) δεν διαφέρουν σημαντικά μεταξύ τους. Μέσες τιμές χωρίς κοινό σύμβολο διαφέρουν σημαντικά μεταξύ τους.

- e. Εκτελέστηκε η ανάλυση διακύμανσης ενός παράγοντα με εξαρτημένη ποσοτική μεταβλητή την αντιλαμβανόμενη ευτυχία (με τιμές 1 έως 10) και παράγοντα την οικογενειακή κατάσταση σε 4 επίπεδα (Γάμος, Διαζύγιο, Χήρος/α, Ποτέ Γάμος) για άτομα άνω των 30 ετών. Οι προϋποθέσεις εκτέλεσης (κανονικότητα και ισότητα διακυμάνσεων) ικανοποιούνται για την εκτέλεση της τυπικής Ανάλυσης Διακύμανσης. Από το αποτέλεσμα της ανάλυσης προέκυψε σημαντική διαφοροποίηση μεταξύ των μέσων τιμών ευτυχίας των τεσσάρων επιπέδων της οικογενειακής Κατάστασης ($F(3,2086)=25,62, p<0,001$); Από τις ανά δύο συγκρίσεις μεταξύ των επιπέδων με τον έλεγχο Bonferroni (σε επίπεδο σημαντικότητας 0,05), διαπιστώνεται ότι οι άνθρωποι άνω των 30 που συμβιώνουν με κάποιο δηλώνουν σημαντικά υψηλότερο επίπεδο ευτυχίας σε σχέση με τους υπόλοιπους στην ίδια ηλικιακή ομάδα (εδώ μπορεί να αναφερθούν οι μέσες τιμές των κατηγοριών). Η κατηγορία των ανθρώπων σε χηρεία δηλώνει το χαμηλότερο επίπεδο ευτυχίας σε σχέση με όλες τις υπόλοιπες κατηγορίες και διαφέρει σημαντικά από την κατηγορία εκείνων που δεν παντρεύτηκαν ποτέ.

3. Θα χρησιμοποιηθεί το αρχείο SATS_PANTEION.sav προκειμένου να ερευνηθεί η υπόθεση ότι οι στάσεις των φοιτητών απέναντι στη στατιστική συνδέονται με τον τύπο πτυχίου στο οποίο αποβλέπουν (μεταβλητή diploma). Θα χρησιμοποιηθούν οι

δύο συνιστώσες των στάσεων C «γνωστικός τομέας» και V «αξία της στατιστικής» που δημιουργήθηκαν στο 2^ο θέμα. Να απαντηθούν τα παρακάτω:

- a. Να γίνουν τα κατάλληλα γραφήματα που περιγράφουν τη σχέση της μεταβλητής diploma με τις στάσεις C και V. Γράψτε συνοπτικά το συμπέρασμά σας.
- b. Ποια είναι η μηδενική υπόθεση στην ανάλυση διακύμανσης και ποια η εναλλακτική;
- c. Προκειμένου να εκτελεστεί η ανάλυση διακύμανσης (one way ANOVA) ποιες προϋποθέσεις πρέπει να ικανοποιούνται;
- d. Να γίνει ο έλεγχος των προϋποθέσεων για την ανάλυση διακύμανσης (one way ANOVA) με εξαρτημένη μεταβλητή την V και της ανάλυσης διακύμανσης με εξαρτημένη την C. Ποια είναι τα συμπεράσματά σας;
- e. Να εκτελεστούν οι κατάλληλοι έλεγχοι (F test ή Welch test) για τις δύο αναλύσεις που αναφέρονται στο παραπάνω ερώτημα και να γραφεί το συμπέρασμά σας σχετικά με την απόρριψη ή μη της μηδενικής υπόθεσης. Να τεκμηριώστε το συμπέρασμά σας δίνοντας απαραίτητα στοιχεία από τους πίνακες του SPSS.
- f. Στο βαθμό που η μηδενική υπόθεση σε κάποια ανάλυση απορρίπτεται στο ερώτημα e, να γίνει ο κατάλληλος έλεγχος προκειμένου να διερευνηθεί ποιος τύπος πτυχίου διαφέρει από ποιον και γραφούν τα συμπεράσματά σας.